

ctc cca caa ggt ata agt tcg ttt att gat gat att cag gct ggt aac 432  
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
130 135 140

tac tcc aca tct tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt 480  
 Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
 145 150 155 160

aaa caa ctc tta cgt gat tat aat atc gag gcc aaa cat cct gtt gta 528  
 Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
 165 170 175

atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att 576  
 Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
 180 185 190

gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg 624  
 Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp  
 195 200 205

gga agt ttt tac atg ctg aga aca ~~aga~~ gtt atg gat aaa gtt tgt tgg 672  
 Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp  
 210 215 220

ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac 720  
 Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn  
 225 230 235 240

ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc 768  
 Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile  
 245 250 255

gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att 816  
 Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
 260 265 270

ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt 864  
 Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
 275 280 285

gca tat tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag 912  
 Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  
 290 295 300

gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta 960  
 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu  
 305 310 315 320

att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg 1008  
 Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp  
 325 330 335

00000-0122500

gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt	1056
Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val	
340 345 350	
aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc	1104
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly	
355 360 365	
gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc	1152
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr	
370 375 380	
att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca	1200
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser	
385 390 395 400	
aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca	1248
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser	
405 410 415	
atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct	1296
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser	
420 425 430	
tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att	1344
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile	
435 440 445	
cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg	1392
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met	
450 455 460	
aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa	1440
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln	
465 470 475 480	
aga aga gta cat gag cag tac tcg ttc ggc tat toc aag aat gaa gaa	1488
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu	
485 490 495	
gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg	1536
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met	
500 505 510	
gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac	1584
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr	
515 520 525	

ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat	1632
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp	
530 535 540	
gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct	1680
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro	
545 550 555 560	
gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca	1728
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser	
565 570 575	
atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga	1776
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly	
580 585 590	
att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat	1824
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp	
595 600 605	
ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc	1872
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser	
610 615 620	
gcg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat	1920
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp	
625 630 635 640	
ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag	1968
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln	
645 650 655	
atg ccc ttc cca atg taa	1986
Met Pro Phe Pro Met	
660	

<213> *Saccharomyces cerevisiae*

225

Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  
340 345 350  
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly  
355 360 365  
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr  
370 375 380  
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser  
385 390 395 400  
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser  
405 410 415  
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser  
420 425 430  
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile  
435 440 445  
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met  
450 455 460  
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  
465 470 475 480  
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu  
485 490 495  
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met  
500 505 510  
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  
515 520 525  
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  
530 535 540  
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro  
545 550 555 560  
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser  
565 570 575  
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  
580 585 590  
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
595 600 605  
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
610 615 620  
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
625 630 635 640  
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
645 650 655  
Met Pro Phe Pro Met  
660

000000-0722560

<210> 3  
<211> 2312  
<212> genomic DNA  
<213> Schizosaccharomyces pombe

<400> 3

ATGGCGTCTT	CCAAGAAGAG	CAAAACTCAT	AAGAAAAAGA	AAGAAGTCAA	50
ATCTCCTATC	GACTTACCAA	ATTCAAAGAA	ACCAACTCGC	GCTTTGAGTG	100
AGCAACCTTC	AGCGTCCGAA	ACACAATCTG	TTTCAAATAA	ATCAAGAAAA	150
TCTAAATTTG	GAAAAAGATT	GAATTTTATA	TTGGGCGCTA	TTTTGGGAAT	200
ATGCGGTGCT	TTTTTTTCG	CTGTTGGAGA	CGACAATGCT	TTTTTCGACC	250
CTGCTACGTT	AGATAAATTT	GGGAATATGC	TAGGCTCTTC	AGACTTGTTT	300
GATGACATTA	AAGGATATTT	ATCTTATAAT	GTGTTTAAGG	ATGCACCTTT	350
TACTACGGAC	AAGCCTTCGC	AGTCTCTAG	CGGAAATGAA	GTTCAAGTTG	400
GTCTTGATAT	GTACAATGAG	GGATATCGAA	GTGACCATCC	TGTTATTATG	450
GTTCTGCTTG	TTATCAGCTC	AGGATTAGAA	AGTTGGTCGT	TTAATAATTG	500
CTCGATTCCCT	TACTTTAGGA	AACGTCTTTG	GGGTAGCTGG	TCTATGCTGA	550
AGGCAATGTT	CCTTGACAAG	CAATGCTGGC	TTGAACATTT	AATGCTTGAT	600
AAAAAAACCG	GCTTGGATCC	GAAGGGAATT	AAGCTGCGAG	CAGCTCAGGG	650
GTTTGAAGCA	GCTGATTTTT	TTATCACCGG	CTATTGGATT	TGGAGTAAAG	700
TAATTGAAAA	CCTTGCTGCA	ATTGGTTATG	AGCCTAATAA	CATGTTAAGT	750
GCTTCTTACG	ATTGGCGGTT	ATCATATGCA	AATTTAGAGG	AACGTGATAA	800
ATATTTTTCA	AAGTTAAAAA	TGTTCAATTGA	GTACAGCAAC	ATTGTACATA	850
AGAAAAAGGT	AGTGTTGATT	TCTCACTCCA	TGGGTTTACA	GGTTACGTAC	900
TATTTTTTTA	AGTGGGTTGA	AGCTGAGGGC	TACGGAAATG	GTGGACCGAC	950
TTGGGTTAAT	GATCATATTG	AAGCATTTAT	AAATGTGAGT	CTCGATGGTT	1000
GTTTGACTAC	GTTTCTAACT	TTTGAATAGA	TATCGGGATC	TTTGATTGGA	1050
GCACCCAAAA	CAGTGGCAGC	GCTTTTATCG	GGTGAAATGA	AAGATACAGG	1100
TATTGTAAAT	ACATTAAACA	TGTTAATATT	TAAATTTTGC	TAACCGTTTT	1150
AAGCTCAATT	GAATCAGTTT	TCGGTCTATG	GGTAAGCAAT	AAATTGTTGA	1200
GATTTGTTAC	TAATTTACTG	TTAGTTTGG	AAAAATTTTT	TTCCCGTTCT	1250
GAGGTATATT	CAAAAATACA	AATGTGCTCT	ACTTTTTCTA	ACTTTTAATA	1300
GAGAGCCATG	ATGGTTTCGA	CTATGGGAGG	AGTTAGTTCT	ATGCTTCCTA	1350
AAGGAGGCGA	TGTTGTATGG	GGAAATGCCA	GTTGGGTAAAG	AAATATGTGC	1400
TGTTAATTTT	TTATTAATAT	TTAGGCTCCA	GATGATCTTA	ATCAAACAAA	1450
TTTTTCCAAT	GGTGCAATTA	TTCGATATAG	AGAAGACATT	GATAAGGACC	1500
ACGATGAATT	TGACATAGAT	GATGCATTAC	AATTTTAAAA	AAATGTTACA	1550
GATGACGATT	TTAAAGTCAT	GCTAGCGAAA	AATTATTCCC	ACGGTCTTGC	1600
TTGGACTGAA	AAAGAAGTGT	TAAAAAATAA	CGAAATGCCG	TCTAAATGGA	1650
TAAATCCGCT	AGAAGTAAGA	ACATTAAAGT	TACTAAATTA	TACTAACCCA	1700
AATAGACTAG	TCTTCCTTAT	GCTCCTGATA	TGAAAATTTA	TTGCGTTCAC	1750
GGGGTCGGAA	AACCAACTGA	GAGAGGTTAT	TATTATACTA	ATAATCCTGA	1800
GGGGCAACCT	GTCATTGATT	CCTCGGTTAA	TGATGGAACA	AAAGTTGAAA	1850
ATGTGAGAGA	ATTTATGTTT	CAAACATTCT	ATTAAGTGTG	TTATTAGGGT	1900
ATTGTTATGG	ATGATGGTGA	TGGAACCTTA	CCAATATTAG	CCCTTGTTTT	1950
GGTGTGCAAT	AAAGTTTGGC	AAACAAAAAG	GTTTAATCCT	GCTAATACAA	2000
GTATCACAAA	TTATGAAATC	AAGCATGAAC	CTGCTGCGTT	TCATCTGAGA	2050
GGAGGACCTC	GCTCGGCAGA	ACACGTCGAT	ATACTTGGAC	ATTCAGAGCT	2100
AAATGTATGT	TCATTTTACC	TTACAAATTT	CTATTACTAA	CTCTTGAAAT	2150
AAGGAAATTA	TTTTTAAAGT	TTCATCAGGC	CATGGTGACT	CGGTACCAAA	2200
CCGTTATATA	TCAGATATCC	AGTACGGACA	TAAGTTTTGT	AGATTGCAAT	2250
TAACATACTA	ACCGAACAGG	GAAATAATAA	ATGAGATAAA	TCTCGATAAA	2300
CCTAGAAATT	AA				

2312

00537740.033000

<210> 4  
 <211> 3685  
 <212> genomic DNA  
 <213> Arabidopsis thaliana

<400> 4

ATGCCCCCTTA	TTCATCGGAA	AAAGCCGACG	GAGAAACCAT	CGACGCCGCC	50
ATCTGAAGAG	GTGGTGCACG	ATGAGGATTC	GCAAAAGAAA	CCACACGAAT	100
CTTCCAAATC	CCACCATAAG	AAATCGAACG	GAGGAGGGAA	GTGGTCGTGC	150
ATCGATTCTT	GTGTTGGGTT	CATTGGGTGT	GTGTGTGTAA	CCTGGTGGTT	200
TCTTCTCTTC	CTTTACAACG	CAATGCCTGC	GAGCTTCCCT	CAGTATGTAA	250
CGGAGCGAAT	CACGGGTCCCT	TTGCCTGACC	CGCCCGGTGT	TAAGCTCAAA	300
AAAGAAGGTC	TTAAGGCGAA	ACATCCTGTT	GTCTTCATTC	CTGGGATTGT	350
CACCGGTGGG	CTCGAGCTTT	GGGAAGGCAA	ACAATGCGCT	GATGGTTTAT	400
TTAGAAAACG	TTTGTGGGGT	GGAACCTTTG	GTGAAGTCTA	CAAAAGGTGA	450
GCTCAACAAT	TCTCACTCTT	CCTTTATATT	GGGATTTGGA	TTGGATCTGA	500
TGAGATCACG	CACTTGTTGC	TCTTCAACA	TCACCTAAAC	TTTAATTCCA	550
TGTTTGTCTG	TCTTACTCTT	TACTTTTTTT	TTTTTTTGAT	GTGAAACGCT	600
ATTTTCTTAA	GAGACTATTT	CTGTATGTGT	AAGGTAAGCG	TTCCAAGGAC	650
GTAATTGGCT	TGGACTATTT	CTGTTTGATT	GTTAACTTTA	GGATATAAAA	700
TAGCTGCCTT	GGAATTTCAA	GTCACTTTAT	TGCCAAATCT	GTTGCTAGAC	750
ATGCCCTAGA	GTCCGTTCAT	AACAAGTTAC	TTCTTTTACT	GTCGTTGCGT	800
GTAGATTAG	CTTTGTGTAG	CGTATATGA	AGTAGTGTTT	TATGTTTTGT	850
TGGGAATAGA	GAAGTTCTAA	CTACATCTGT	GGAAAGTGTG	TTCAGGCTGT	900
GATAGAGGAC	TGTTGCTTTA	TTATTCAACT	ATGTATATGT	GTAATTAAAG	950
CTAGTTCCTT	TTTGATCTTT	CAGCTCAATT	TGCTTTTCTC	AATTTTTTTC	1000
TCAATTTCAA	AGTTTCACAT	CGAGTTTATT	CACATGTCTT	GAATTTTCGT	1050
CATCCTCGTT	CTGTTATCCA	GCTTTGAACT	CCTCCCGACC	CTGCTATGGA	1100
TATATTAAAA	AAAAAGTGTT	TTGTGGGTTG	CATCTTTGTT	ACGATCTGCA	1150
TCTTCTTCTT	TCGGCTCAGT	GTTTCATGTT	TTGCTATGGT	AGAGATGGGC	1200
AATGTTATTG	TTGATGGTAA	CAGTGGTATA	GTTGATAGTA	TCTTAACTAA	1250
TCAATTATCT	CTTTGATTCA	GGCCTCTATG	TTGGGTGGAA	CACATGTCAC	1300
TTGACAATGA	AACTGGGTTG	GATCCAGCTG	GTATTAGAGT	TCGAGCTGTA	1350
TCAGGACTCG	TGGCTGCTGA	CTACTTTGCT	CCTGGCTACT	TTGTCTGGGC	1400
AGTGCTGATT	GCTAACCTTG	CACATATTGG	ATATGAAGAG	AAAAATATGT	1450
ACATGGCTGC	ATATGACTGG	CGGCTTTCGT	TTCAGAACAC	AGAGGTTCTT	1500
TTCTCATCGT	TCTTCTTATT	ATTCTGTTCC	ATGTTACGTT	TCTTCTTCA	1550
TTACTTAAGG	CTTAAATATG	TTTCATGTTG	AATTAATAGG	TACGTGATCA	1600
GACTCTTAGC	CGTATGAAAA	GTAATATAGA	GTTGATGGTT	TCTACCAACG	1650
GTGGAAAAAA	AGCAGTTATA	GTTCCGCATT	CCATGGGGGT	CTTGTATTTT	1700
CTACATTTTA	TGAAGTGGGT	TGAGGCACCA	GCTCCTCTGG	GTGGCGGGGG	1750
TGGGCCAGAT	TGGTGTGCAA	AGTATATTAA	GGCGGTGATG	AACATTGGTG	1800
GACCATTCTT	TGGTGTTCCT	AAAGCTGTTG	CAGGGCTTTT	CTCTGCTGAA	1850
GCAAAGGATG	TTGCAGTTGC	CAGGTATTGA	ATATCTGCTT	ATACTTTTGA	1900
TGATCAGAAC	CTTGGCTCTG	GAACTCAAAG	TTATCTTACT	AAATATCAAT	1950
TCTAATAACA	TTGCTATATT	ATCGCTGCAA	CTGACATTGG	TTGATTATTT	2000
TTGCTGCTTA	TGTAAGTGAA	ACTCTCTTGA	GATTAGACAA	ATGATGAATT	2050
GATAATTCTT	ACGCATTGCT	CTGTGATGAC	CAGTTTCTTA	GCTTCGACGA	2100
TAACATTTGT	CATACTGTCT	TTTGGAGGGC	ATTGAATTTT	GCTATGGAAA	2150
GCGCTGGAGC	TTCCATGCTT	GCATTCTTTA	CCAATTAGCG	TTATTCTGCT	2200
TCTTTCAAAT	TTCTTGTATA	TGCATCTATG	GTCTTTTATT	TCTTCTTAAT	2250
TAAAGACTCG	TTGGATTAGT	TGCTCTATTA	GTCACCTGGT	TCCTTAATAT	2300
AGAACCTTAC	TTTCTTCGAA	AATTGCAGAG	CGATTGCCCC	AGGATTCTTA	2350
GACACCGATA	TATTTAGACT	TCAGACCCTG	CAGCATGTAA	TGAGAAATGAC	2400
ACGCACATGG	GACTCAACAA	TGTCTATGTT	ACCGAAGGGA	GGTGACACGA	2450
TATGGGGCGG	GCTTGATTGG	TCACCGGAGA	AAGGCCACAC	CTGTTGTGGG	2500
AAAAAGCAAA	AGAACAACGA	AACTTGTGGT	GAAGCAGGTG	AAAACGGAGT	2550
TTCCAAGAAA	AGTCCTGTTA	ACTATGGAAG	GATGATATCT	TTTGGGAAAG	2600
AAGTAGCAGA	GGCTGCGCCA	TCTGAGATTA	ATAATATTGA	TTTTCGAGTA	2650
AGGACATATA	AATCATAATA	AACCTTGATC	ATTTGTGTAT	TGTATGATGA	2700
ATATCTGTAC	ATTTTATCTG	GTGAAGGGTG	CTGTCAAAGG	TCAGAGTATC	2750
CCAAATCACA	CCTGTCTGTA	CGTGTGGACA	GAGTACCATG	ACATGGGAAT	2800
TGCTGGGATC	AAAGCTATCG	CTGAGTATAA	GGTCTACACT	GCTGGTGAAG	2850

00537710.033000



CTATAGATCT	ACTACATTAT	GTTGCTCCTA	AGATGATGGC	GCGTGGTGCC	2900
GCTCATTCT	CTTATGGAAT	TGCTGATGAT	TTGGATGACA	CCAAGTATCA	2950
AGATCCCAAA	TACTGGTCAA	ATCCGTTAGA	GACAAAGTAA	GTGATTTCTT	3000
GATTCCAAC	GTATCCTTCG	TCCTGATGCA	TTATCAGTCT	TTTTGTTTTT	3050
GGTCTTGTTG	GATATGGTTT	TCAGCTCAAA	GCTTACAAAG	CTGTTTCTGA	3100
GCCTTTCTCA	AAAAGGCTTG	CTCAGTAATA	TTGAGGTGCT	AAAGTTGATA	3150
CATGTGACTC	TTGCTTATAA	ATCCTCCGTT	TGGTTTGTTT	TGCTTTTTTCA	3200
GATTACCGAA	TGCTCCTGAG	ATGGAATGAT	ACTCATTATA	CGGAGTGGGG	3250
ATACCAACGG	AACGAGCATA	CGTATACAA	CTTAACCAGT	CTCCCGACAG	3300
TTGCATCCCC	TTTCAGATAT	TCACTTCTGC	TCACGAGGAG	GACGAAGATA	3350
GCTGTCTGAA	AGCAGGAGTT	TACAATGTGG	ATGGGGATGA	AACAGTACCC	3400
GTCCCTAAGTG	CCGGGTACAT	GTGTGCAAAA	GCGTGGCGTG	GCAAGACAAG	3450
ATTCAACCTT	TCCGGAATCA	AGACTTATAT	AAGAGAATAC	AATCACTCTC	3500
CGCCGGCTAA	CCTGTTGGAA	GGGCGCGGGA	CGGAGAGTGG	TGCCCATGTT	3550
GATATCATGG	GAAACTTTGC	TTTGATCGAA	GATATCATGA	GGGTTGCCCG	3600
CGGAGGTAAC	GGGTCTGATA	TAGGACATGA	CCAGGTCAC	TCTGGCATAT	3650
TTGAATGGTC	GGAGCGTATT	GACCTGAAGC	TGTGA		

[illegible]

```
<210> 5
<211> 2427
<212> cDNA
<213> Arabidopsis thaliana
```

Sequence	Count
AGAAACAGCT	50
TTCTAAATTC	100
TTAATTTCAA	150
AGAAACCATC	200
CAAAAGAAAC	250
AGGAGGGAAG	300
TGTGTGTAAC	350
AGCTTCCCTC	400
GCCCGGTGTT	450
GTCTTCATTC	500
ACAATGCGCT	550
GTGAAGTCTA	600
AATGAAACTG	650
ACTCGTGGCT	700
TGATTGCTAA	750
GCTGCATATG	800
GA CTCTTAGC	850
GTGGAAGAAA	900
CTACATTTTA	950
TGGGCCAGAT	1000
GACCATTTCT	1050
GCAAAGGATG	1100
CGATATATTT	1150
CATGGGACTC	1200
GGCGGGCTTG	1250
GCAAAGAAGC	1300
AGAAAAGTCC	1350
GCAGAGGCTG	1400
CAAAGGTCAG	1450
ACCATGACAT	1500
TACATGCTCG	1550
GATGGCGCGT	1600
ATGACACCAA	1650
AAATTACCGA	1700
GATACCAACG	1750
GTTGCATCCC	1800
AGCTGTCTGA	1850
CGTCCTAAGT	1900
GATTCAACCC	1950
CCGCCGCGTA	2000
TGATATCATG	2050
CCGGAGGTAA	2100
TTTGAATGGT	2150
TTAAGCTGTC	2200
ATCATCAATT	2250
AGCCTGAGAA	2300
CTTATTGAAT	2350
AAAATTTGCT	2400
GGGTTAAAAA	

<210> 6  
<211> 671  
<212> PRT  
<213> Arabidopsis thaliana

<400> 6  
MPLIHRKKPT EKPSTPPSEE VVHDEDSQKK PHESSKSHHK KSNNGGKWSC 50  
IDSCCWFIGC VCVTWWFLLF LYNAMPASFP QYVTERITGP LPDPPGVKLK 100  
KEGLKAKHPV VFIPGIVTGG LELWEGKQQA DGLFRKRLWG GTFGEVYKRP 150  
LCWVEHMSLD NETGLDPAGI RVRVSGLEVA ADYFAPGYFV WAVLIANLAH 200  
IGYEEKNMYM AAYDWRLSFQ NTEVRDQTLS RMKSNIELMV STNGGKKAVI 250  
VPHSMGVLYF LHFMKWVEAP APLGGGGGPD WCAKYIKAVM NIGGPFLGVP 300  
KAVAGLFSAE AKDVAVARAI APGFLDTDIF RLQTLQHV MR MTRTWDSTMS 350  
MLPKGGDTIW GGLDWSPEKG HTCCGKKQKN NETCGEAGEN GVSCKSPVNY 400  
GRMISFGKEV AEAAPSEINN IDFRGAVKGQ SIPNHTCRDV WTEYHDMGIA 450  
GIKAIAEYKV YTAGEAIDLL HYVAPKMMAR GAAHFSYGIA DDLDDTKYQD 500  
PKYWSNPLET KLPNAPEMEI YSLYGVGIPT ERAYVYKLNQ SPDSCIPFQI 550  
FTSAHEDED SCLKAGVYNV DGDETVPVLS AGYMCACAWR GKTRFNPSGI 600  
KTYIREYNHS PPNALLEGRG TQSGAHVDIM GNFAIEDIM RVAAGNGSD 650  
IGHDQVHSGI FEWSEIDLK L 671

00000-01228560

```
<221> CDS
<222> (1)..(402)
```

<400> 6  
CGG GAG AAA ATA GCT GCT TTC AAG GGG GGT GTT TAC TTA GCC GAT GGT 48  
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly  
1 5 10 15

GAT GAA ACT GTT CCA GTT CTT AGT GCG GGC TAC ATG TGT GCG AAA GGA 96  
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly  
20 25 30

TGG CGT GGC AAA ACT CGT TTC AGC CCT GCC GGC AGC AAG ACT TAC GTG 144  
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val  
35 40 45

AGA GAA TAC AGC CAT TCG CCA CCC TCT ACT CTC CTG GAA GGC AGG GGC 192  
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly  
50 55 60

ACC CAG AGC GGT GCA CAT GTT GAT ATA ATG GGG AAC TTT GCT CTA ATT 240  
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile  
65 70 75 80

GAG GAC GTC ATC AGA ATA GCT GCT GGG GCA ACC GGT GAG GAA ATT GGT 288  
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly  
85 90 95

GGC GAT CAG GTT TAT TCA GAT ATA TTC AAG TGG TCA GAG AAA ATC AAA 336  
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys  
100 105 110

TTG AAA TTG TAA CCTATGGGAA GTTAAAGAAG TGCCGACCCG TTTATTGCGTTCC 391  
Leu Lys Leu  
115

AAAGTGTCTT	GCCTGAGTGC	AACTCTGGAT	TTTGCTTAAA	TATTGTAATT	TTTCACGC	449
TTCATTCTGTC	CCTTTGTCAA	ATTTACATTT	GACAGGACGC	CAATGCGATA	CGATGTTG	507
TACCGCTATT	TTTCAGCATTG	TATATTAAAC	TGTACAGGTG	TAAGTTGCAT	TTGCCAGC	565
TGAAATTGTG	TAGTCGTTTT	CTTTACGATT	TAATANCAAG	TGGCGGAGCA	GTGCCCCA	623
AGCNAAAAAA	AAAAAAAAAA		643			

```

<400> 8
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly
  1                               10                      15

Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly
      20                               25                      30

Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val
      35                               40                      45

Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly
      50                               55                      60

Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile
  65                               70                      75                      80

Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly
      85                               90                      95

Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Ly
      100                               105                      110

Leu Lys Leu
      115

```

<400> 9'

<400>	ggtggcggaag	acgangtcgcg	aagttggagg	ctaacgagaa	tgacnctcgg	50
	agatggatct	acctctgaga	gacacgagta	cnnttgacc	cagcctcaag	100
	gtntacngtt	tntatgggta	ggaagcggac	ggagcgagcc	tacatctatc	150
	tggcgccga	tcccgggacg	acaacgcac	tttagatgac	gatcgatacg	200
	actttgactn	aggggcacat	tgaccacggt	gtgattttgg	gcgaaggcga	250
	tggcacagtg	aaccttatga	gtttggggta	cctgtgcaat	aaggggtgga	300
	aaatgaagag	atacaatcct	gcgggctcaa	aaataaccgt	ggtcgagatg	350
	ccgcatgaac	cagaacggtt	caatccgaga	ggagggccga	atacggcgga	400
	tcacgtggat	attctaggaa	ggcagaatct	aaacgagtac	attcttaaag	450
	tggcggcagg	tcgaggcagt	acaattgagg	attttattac	tagtaatat	500
	cttaaatatg	tagaaaaggt	tgaattttat	gaagagtaat	taaatacggc	550
	acataaggtta	ctcaatagta	tgactaatta	aaaaaaaaat	ttttttctaa	600
	aaaaaaaaaa	aaaaaa				

[illegible]

<400> 10

2400> 10	50	
ATGAAAAAAA	TATCTTACAA	TTATTCCGTA
GGTGACGATG	ACCTCGATGT	GTCAAGCTGT
TGATTCTGGT	TCCAGGAAAC	GGAGGTAACC
AGAGAATACA	AGCCAAGTAG	TGTCTGCTGT
TCATAAGAAG	AGTGGTGGAT	GGTTTAGGCT
TATTGTCTCC	CTTCACCAGG	TGCTTCAGCG
GACCCTGATT	TGGATGATTA	CCAAAATGCT
TCTCATTTTC	GGTTCGACCA	AATCACTTCT
GGTTAGTACT	TTCCAAGATA	TATCATTTTG
AAATAGACAT	AAATTTGGGG	GATTATTGTT
GCTAGTCGGT	AATGTGAGTG	TTATGTTAGT
GTGATTTTCC	ATTTTAAATG	AAGCTAGAAA
CTATGTCATG	AGAATTATAA	GGACACTATG
GGTTTGATTT	GCAGAGATGC	CACATCTTAC
TCTAGAGAAA	AAATGCGGGT	ATGTTAACGA
CATATGATTT	CAGGTACGGC	CTGGCTGCCT
GCCTCACAGT	TCCTACAAGA	CCTCAAACAA
CGAGAACGAA	GGAAAGCCAG	TGATACTCCT
TTTTCGTCCT	CCATTTCCCTC	AACCGTACCA
TACATCAAAC	ACTTTGTTGC	ACTCGCTGCG
TCAGATGAAG	ACATTTGCTT	CTGGCAACAC
ACCCTTTGCT	GGTCAGACGG	CATCAGAGGA
CTACTTCCAT	CTACCAAAGT	GTTTCACGAC
AACTCCCCAG	GTAACTACA	CAGCTTACGA
ACATTGGATT	CTCACAAGGA	GTTGTGCCTT
TTAACAGAGG	AGCTGATGAC	TCCGGGAGTG
GAGAGGAGTT	GATACACCGG	AGGTTTTGAT
ATAAGCAACC	AGAGATTAAG	TATGGAGATG
GCGAGCTTAG	CAGCTTTTGA	AGTCGATAGC
TGGAGTTTCG	CATCATCTCTA	TACTTAAAGA
TTATGAAGCA	GATTTCAATT	ATTAATTATG
GTCAATGAAT	GA	

50  
100  
150  
200  
250  
300  
350  
400  
450  
500  
550  
600  
650  
700  
750  
800  
850  
900  
950  
1000  
1050  
1100  
1150  
1200  
1250  
1300  
1350  
1400  
1450  
1500  
1550  
1600  
1650  
1700  
1750  
1800  
1850  
1900  
1950  
2000  
2050  
2100  
2150  
2200  
2250  
2300  
2350  
2400  
2450  
2500  
2550  
2600  
2650  
2700  
2750  
2800  
2850

ATGGGAGCGA ATTCGAAATC AGTAACCGCT TCCTTCACCG TCATCGCCGT  
 TTTTTCTTG ATTTGCGGTG GCGGAACTGC GGTGGAGGAT GAGACCGAGT  
 TTCACGGCGA CTA CTCTCGAAG CTATCGGGTA TAATCATTCC GGGATTTGCG  
 TCGACGCAGC TACGAGCGTG GTCGATCCTT GACTGTCCAT AACTCCTGTT  
 GGACTTCAAT CCGCTCGACC TCGTATGGCT AGACACCACT AAGGTCCGTG  
 ATCTTCATTT CCTTCGCTCC TTATCTGTGTC GGTCGAGTCA CTTGTTGATG  
 AATTCCAAGC GAAATATAGC AATGAAGCAT GTCTCGTCTC TCTTATTGAT  
 TCGTTTCATTA GTCAACAGTG ACGCTTCTGA ATCTGAGTTT AGAGTCATAT  
 AAAACAGCTG ACTCGGCGAG TGTTCCTCAT CGCTTTTGGT TCGCTAAATG  
 TAGCGCAATG AATGTGTAAT TAGTCTCCGC TTTTTATTCA ACTAGATCTG  
 CAAGTTTCTT AGATGCTCTCA ATAGTAGTTA GAAAATGTTA GGTCAATTTA  
 CTTGTGCAAT GTGATTCTTT TGGTTGTGTC TTACTGATCG ACGTGATGGA  
 TGGTTTACAG CTTCTTTCTG CTGTCAACTG CTGGTTTAAG TGTATGGTGC  
 TAGATCCTTA TAATCAAACA GACCATCCCG AGTGTAAGTC ACGGCTGAC  
 AGTGGTCTTT CAGCCATCAC AGAATTGGAT CCAGGTTACA TAACAGGTAG  
 TTTCGGATTT TTCTTTCTTT TGAGTTTTCT TCAATTTGAT ATCATCTTGT  
 TGTGATATAA TATGGCTAAG TTCATTAATT TGGTCAATTT TCAGGTCCCTC  
 TTTCTACTGT CTGGAAAGAG TGGCTTAAAG GGTGTGTTGA GTTTGGTATA  
 GAAACAAATG CAATGTGCGC TGTCCATAC GATTGGAGAT GTCCACCAAC  
 CAAATTTGAA GAGCGTGACC TTTACTTTCA CAGCTCAAG TTAGTCTTCA  
 TCAGGCTAAT GTCTTTTATC TTCTCTTTTT ATGTAAGATA AGCTAAGAGC  
 TCTGGTCGTC TTCCTTTTTG CAGGTTGACC TTTGAAACTG CTTTAAACT  
 CCGTGGCGGC CCTTCTATAG TATTTGCCCA TTCATGGGT AATAATGTCT  
 TCAGATACTT TCTGGAATGG CTGAGGCTAG AAATGACACC AAAACATTAT  
 TTGAAGTGGC TTGATCAGCA TATCCATGCT TATTTGCTG TTTGTTACCGG  
 CCTACTATCC TTAAGTTACC ATTTTATTT TTCTCTAATT GGGGGAGTTA  
 TGTGTGAGT TACTGGATTG AGCTCGATAC CTGATTGTGT GTTGATTAG  
 GAGCTCCTCT TCTTGGTTCT GTTGAGGCAA TCAAATCTAC TCTCTCTGGT  
 TTAACGTTTG GCCTTCTGT TTCTGAGGTG ACCTCTGACT TCTCTTTAGT  
 TGAAAGTATT ACTTTTGTTA ATTGAACTGC AACTCACTGG ATTTTCTTT  
 TAGATCTTGA AGTGCTAGTT ATCAAAGAAC TGTACGCGAT ATGGTATCTG  
 GTCAGCGGCC TTAGCTAATA CAACCAAACC ACATGTACAC TGATTAGTT  
 TTCAGATTAT TATGGTAGAC TTTAAGTTGA GAAGAAACTT TGACTGAAAT  
 CTTTTTATTT TAATAGGCTA TGATTGTGTT ATTGAAATCA TGTGACATAT  
 TGACATGCGC TTCTCATGTT TTTTGTGGC AAGGCTTCAG GGAAGTCTC  
 GGTGTGTTGC CAATCTTTT CCGTCTCAT TGTGGCTTAT GCCATTTTCA  
 AAGAATTGCA AGGGTGATAA CACATTCTGG ACGCATTTTT CTGGGGGTGC  
 TGCAAAGAAA GATAAGCGCG TATACCACTG TGATGAAGAG GAATATCAAT  
 CAAAATATTC TGGCTGGCCG ACAAATATTA TTAACATTGA AATTCCTTCC  
 ACTAGCGGTT AGACTCTGTA TATGCAACTG TAACACTAAC AAAAGTTTCA  
 CCAAGAATGT TCACTCTCAT ATTTCTGTTCC TTTGATGTGT ATCCATCAGT  
 TACAGAAACA GCTCTAGTCA ACATGACCAG CATGGAATGT GGCCTTCCCA  
 CCCTTTTGTC TTTCACAGCC AGTGAACAGT CAGATGGGAC TCTTTTCAAA  
 GCAATAGAAG ACTATGACCC CGTAGACTAG AGGATGTTAC ACCAGTTAAA  
 GAAGTACGTA CCTTTCTTTG TGATAAGAAA TATTGCTCAT CGATCATCAC  
 TTGCTGGCTT CTGTGACGTC AAATTGTTTT GTTTAAATCT CTATATCAAT  
 TGTTCATATG CTTTGTCTTT CTTACTATAA GAAACAAGTA TAATCAGAAA  
 CCTTATTATT GATTATCAGT TCTCTCCTTA TATTATGGAA TGTCTTTTTC  
 GTTTACAGTT ATGAATGCAA AAGGGGGTAT TTTAGTTGAT TGATTCTCTC  
 ATTCTCTAGT TTGTTTTGAC TAATAGCGTC AATTTTGTTT TTTAGCAAA  
 TCTTTGTGAA TTATATATAA CATGCTAACT ATACTTTTCA GGTGTATCA  
 TGATGACCCCT GTTTTTAATC CTCTGACTCC TTGGGAGAGA CCACCTATAA  
 AAAATGTATT TTGCATATAT CGTGCTCATC TAAAGACAGA GGTATGATGC  
 ATTCTCAATA TCACATTATG CGTTGACTTT GTTATTATAT TCCCCATTTC  
 GTTTGCAATA TCTTTTTGAA TTATGATTTA TCTTCTCCCT TGCATCTTAT  
 GCTATTAAGC GTTAAAGGTA CTAATGTAT GAAGCTGTCT GTCATAGGTT



GGTTATTACT	TTGCCCAAG	TGGCAAACCT	TATCCTGTGATA	ATTGGATCAT	2900
CACGGATATC	ATTTATGAAA	CTGAAGGTTC	CTCCGTGTCA	AGGTAATTTT	2950
CCGCAATGTC	AGAAGTAAAA	CAGGAAGGCA	AAGTCTTCTG	TATCAGTCTA	3000
GTGGCATGTT	ATCTCAGTTG	GATAAGCAAA	TTATTAAACA	ACTAAAATTT	3050
AAGTACTTTT	TTATCATTCC	TTTTGAGCTT	AGTGGATGAT	CAGTGGCTTA	3100
AAGTGGGAAG	AGGTGTTGCA	TGAAACATGA	CACTTGTATC	AAAGATAACT	3150
AGCAAAACAA	AACATAACCA	TTTCTGAATT	TCATATTATT	AGGAGTAGTC	3200
GTGCTTTTAA	AAAATTTGTT	TTAAGAAACC	GAAAAAGTAG	TTCATATCTT	3250
GATTGTGCAA	TATCTGCAGG	TCTGGTCTCG	TGGTTGATGG	GAAACGCTGA	3300
CCTATAACTG	GGGATGAGAC	GGTAACTCTA	GAAGTTGGTT	TTGAAATTAT	3350
CTTCTTGCAA	ACTACTGAAG	ACTAAGATAA	TACTTGCTTC	TGGAACACTG	3400
CTTGCTATGT	TCTCTAGTAC	ACTGCAATAT	TGACTCTCCG	CTACTTTTAT	3450
TGATTATGAA	ATTGATCTCT	TATAGGTACC	CTATCATTCA	CTCTCTTGGT	3500
GCAAGAATTG	GCTCGGACCT	AAAGTTAACA	TAACAATGGC	TCCCCAGGTA	3550
CTCTTTTTTTA	GTTCCCTCACC	TTATATAGAT	CAAACTTTAA	GTGTACTTTT	3600
CTGGTTATGT	GTTGATTTAC	CTCCAATTTG	TTCTTTCTAA	AAATCATATA	3650
TCTCTGTACT	CCTCAAGAAC	TTGTATTAAT	CTAAACGAGA	TTCTCATTGG	3700
GAAAAATAAAA	CAACAGCCAG	AACACGATGG	AAGCGACGTA	CATGTGGAAC	3750
TAAATGTTGA	TCATGAGCAT	GGTTCAGACA	TCATAGCTAA	CATGACAAAA	3800
GACCAAGGG	TTAAGTACAT	AACCTTTTAT	GAAGACTCTG	AGAGCATTCC	3850
GGGGAAGAGA	ACCGCAGTCT	GGGAGCTTGA	TAAAGTGGG	TATTAA	3896

<400> 12

CTGGGGCCAA	AAGTGAACAT	AACAAGGACA	CCACAGTCAG	AGCATGATGT	50
TCAGATGTAC	AAGTGCATCT	AAATATAGAG	CATCAACATG	GTGAAGATAT	100
CATTCCCAAT	ATGACAAAGT	TACCTACAAT	GAAGTACATA	ACCTATTATG	150
AGGATTCTGA	AAGTTTTCCA	GGGACAAATA	CAGCAGTTTG	GGAGCTTGAT	200
AAAGCAAATC	ACAGGAACAT	TGTCAGATCT	CCAGCTTTGA	TGCGGGAGCT	250
GTGGCTTGAG	ATGTGGCATG	ATATTCAATC	TGATAAAAAG	TCCAAGTTTG	300
TTACAAAAGG	TGGTGTCTGA	TCCTCACTAT	TTTCTTCTAT	AAATGTTTGA	350
GTTTGTATTG	ACATTGTAAG	TATTGCAACA	AAAAGCAAAG	CGTGGGCCTC	400
TGAGGGATGA	GGAGTGCTAT	TGGGATTACG	GGAAAGCTCG	ATGTGCATGG	450
GCTGAACATT	GTGAATACAG	GTTAGAATAT	TCAAATTATA	TTTTGCAAAA	500
TATTCTCTTT	TGTGTATTT	AGGCCACCTT	TCCCCGGTCA	CAACGATGCA	550
GATATGTATT	CGGGGATGTT	CACCTGGGAC	AGAGTTGCAG	ATTGAAGAGT	600
TCTACATCTC	ACATCCTGTC	ACACTATGTG	TGATATTTAA	GAAACTTTGT	650
TTGGCGGAAC	AACAAGTTTG	CACAAACATT	TGAAGAAGAA	AGCGAAATGA	700
TTCAGAGAG					

[illegible]

<400> 13

MASSKSKTHKKKKEVKSPIDL PNSKKPTRL	SEQPSASETQSVSNKSRKSKFGKRLNFI	LGA	LIGICGA	70
FFFAVGDDNAVFD PATLDKFGNMLGSSD	LDLDDIKGYLSYNVFKDAPFTTDKPSQ	SPSGNEVQVGLD	MYNE	140
GYRSDHPVIMVPGVISSGLESWSFNCS	IPYFRKRLWGSWSMLKAMFLDKQCWLE	HLMLDKKTGLD	PKGI	210
KLRAAQGFEEADFFITGYWIWSKVIEN	LAAIEYEPNNMLSASYDWRLSYANLEER	DKYF SKLKMFI	EYSN	280
IVHKKKVVLI SHSMGSQVTTYFFKWEA	EAGYGNCGGPTWVNDHIEAFINISGSL	IGAPKTVAALL	SGEMKD	350
TGIVITLNLIEKFFSRSERAMMVRTMG	GVSSMLPKGGDVAPDDL NQTNF SNGAI	IRYREDIDKD	HDEFDI	420
DDALQFLKNVTDDDFKVMLAKNYSHGL	AWEKEV LKNNEMPSKWINPLETSLPYA	PD MKIYCVH	GVGKPT	490
ERGYYYTNNPEGQPVIDSSVNDGTKVEN	GI VMDGDGTLPIALGLVCNKVWQTKR	FN PANTSIT	NYEIK	560
HEPAAFDLRGGPRSAEHVDILGHSELNE	II LKVSSGHGDSVPNRYISDIQEI	INEINL	DKPRN	623

&lt;210&gt; 14

&lt;211&gt; 432

&lt;212&gt; PRT

<213> *Arabidopsis thaliana*

&lt;400&gt; 14

MKKISSHYSVVIAILVVVTMTSMCQAVGSNVPLILVPGNGGNQLEVRDREYKPSSVWCSSWLYPIHKK 70  
SGGWFRWFDAAVLLSPFTRCFSDRMMLYYDPDLDDYQNA PGVQTRVPHFGSTKSLLYLDPRLRDATSYM 140  
EHLVKALEKKCGYVNDQITLGAPYDFRYGLAASCHPSRVASQFLQDLKQLVEKTSSENEKGPVILLSHSL 210  
GGLFVLHFLNRTTPSWRRKYIKHFVALAAPWG GTISQMKTFASGNTLG VPLVNPLLVR RHQRTSES NQWL 280  
LPSTKVFHDRTKPLVVT PQVNYTAYEMDRFFADIGFSQGVVPYKTRVLP LTEELMTPGVPVTCIYGRGVD 350  
TPEVL MYGKGGFDKQPEIKYGDGDGT VNLASLAALKVDSLNTVEIDGVSHTSILKDEIALKEIMQISII 420  
NYELANVNAVNE

432

00000-0422560

<400> 15

MGANSKSVTASFVTIAVFFLICGGRTAVEDETEFHGDYSKLSGIIIPGFASTQLRAWSILDCPYTPLDFN 70  
PLDLVWLDTTKLLSAVNCWFKCMVLPYNOTDHPECKSRPDSGLSAITELDPGYITGPLSTVWKEWLKWC 140  
VEFGIEANAIVAVPYDWRLSPTKLEEROLVPHKLKLTFTETALKLRGGPSIVFAHSMGNVNFYFLEWLRL 210  
EIAPKHYLKWLDQHIHAYFAVGAPLLGSVEAIKSTLSGVTFGLPVSEGTARLLSNSFASSLWLMPPFSKNC 280  
KGDNTFWTHFSGGAAKKDKRVYHCDEEEYQSKYSGWPTNIINIEIPSTSARELADGTLFKAIEDYDPDSK 350  
RMLHQLKKYVPPFVIRNIAHRSSLAGFLLYHDDPVFNPLTPWERPPIKNVFCIYGAHLKTEVGYYPAPSG 420  
KPYPDNWIITDIIYETEGSLVSRSGTVVDGNAGPITGDETVPYHSLSWCKNLGPKVNITMAPQILIGKI 490  
KQPEHDGSDVHVHVLNVDHEHGSIIANMTKARRVKYITFYEDSESI PGKRTAVWELDKSGY 552

Figure 1 displays 12 histograms, labeled  $x_0$  through  $x_{11}$ , showing the distribution of the number of non-zero elements in the vector  $x_k$ . The x-axis represents the number of non-zero elements (0 to 10), and the y-axis represents the count (0 to 10). The distributions are roughly bell-shaped and centered around 5, with the peak count increasing from 10 at  $x_0$  to 12 at  $x_{11}$ .

10

15

<210> 1a

<211> 661

20

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 1

25

```

<400> 1
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
      1              5              10              15

```

Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
20 25 30

30

Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
35 40 45

Ile Ser Gly Ser Ala Lys Arg ~~Asn~~ Glu Arg Gly Lys Asp Phe Asp Arg  
50 55 60

35

Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
65 70 75 80

40

Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
85 90 95

Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
100 105 110

45

Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
115 120 125

Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
130 135 140

50

Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
145 150 155 160

55

Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
165 170 175

Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
180 185 190

60

Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp  
195 200 205

Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp  
210 215 220

5 Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn  
225 230 235 240

Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile  
245 250 255

10 Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
260 265 270

Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
275 280 285

15 Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  
290 295 300

20 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu  
305 310 315 320

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp  
325 330 335

25 Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  
340 345 350

30 Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly  
355 360 365

Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr  
370 375 380

35 Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser  
385 390 395 400

Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser  
405 410 415

40 Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser  
420 425 430

Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile  
435 440 445

45 Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met  
450 455 460

50 Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  
465 470 475 480

Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu  
485 490 495

55 Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met  
500 505 510

Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  
515 520 525

60

00000-0422500

Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  
 530 535 540  
 5 Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro  
 545 550 555 560  
 Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser  
 565 570 575  
 10 Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  
 580 585 590  
 Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
 595 600 605  
 15 Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
 610 615 620  
 20 Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
 625 630 635 640  
 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
 645 650 655  
 25 Met Pro Phe Pro Met  
 660  
 30 <210> 2a  
 <211> 387  
 <212> PRT  
 <213> Arabidopsis thaliana  
 35 <400> 2  
 Val Gly Ser Asn Val Tyr Pro Leu Ile Leu Val Pro Gly Asn Gly Gly  
 1 5 10 15  
 Asn Gln Leu Glu Val Arg Leu Asp Arg Glu Tyr Lys Pro Ser Ser Val  
 20 25 30  
 40 Trp Cys Ser Ser Trp Leu Tyr Pro Ile His Lys Lys Ser Gly Gly Trp  
 35 40 45  
 45 Phe Arg Leu Trp Phe Asp Ala Ala Val Leu Leu Ser Pro Phe Thr Arg  
 50 55 60  
 Cys Phe Ser Asp Arg Met Met Leu Tyr Tyr Asp Pro Asp Leu Asp Asp  
 65 70 75 80  
 50 Tyr Gln Asn Ala Pro Gly Val Gln Thr Arg Val Pro His Phe Gly Ser  
 85 90 95  
 Thr Lys Ser Leu Leu Tyr Leu Asp Pro Arg Leu Arg Asp Ala Thr Ser  
 100 105 110  
 55 Tyr Met Glu His Leu Val Lys Ala Leu Glu Lys Lys Cys Gly Tyr Val  
 115 120 125  
 60 Asn Asp Gln Thr Ile Leu Gly Ala Pro Tyr Asp Phe Arg Tyr Gly Leu  
 130 135 140



5

10

15

20

24

30

3.

4

4

(7)

12

6

	Asp	Gly	Leu	Phe	Arg	Lys	Arg	Leu	Trp	Gly	Gly	Thr	Phe	Leu	Cys	Trp
	35							40						45		
5	Val	Glu	His	Met	Ser	Leu	Asp	Asn	Glu	Thr	Gly	Leu	Asp	Pro	Ala	Gly
	50						55					60				
	Ile	Arg	Val	Arg	Ala	Val	Ser	Gly	Leu	Val	Ala	Ala	Asp	Tyr	Phe	Ala
	65					70					75				80	
10	Pro	Gly	Tyr	Phe	Val	Trp	Ala	Val	Leu	Ile	Ala	Asn	Leu	Ala	His	Ile
					85					90					95	
	Gly	Tyr	Glu	Glu	Lys	Asn	Met	Tyr	Met	Ala	Ala	Tyr	Asp	Trp	Arg	Leu
				100					105					110		
15	Ser	Phe	Gln	Asn	Thr	Glu	Arg	Asp	Gln	Thr	Leu	Ser	Arg	Met	Lys	Ser
			115					120					125			
	Asn	Ile	Glu	Leu	Met	Val	Ser	Thr	Asn	Gly	Gly	Lys	Lys	Ala	Val	Ile
20			130				135					140				
	Val	Pro	His	Ser	Met	Gly	Val	Leu	Tyr	Phe	Leu	His	Phe	Met	Lys	Trp
	145					150				155						160
25	Val	Glu	Ala	Pro	Ala	Pro	Leu	Gly	Gly	Gly	Gly	Gly	Pro	Asp	Trp	Cys
					165					170					175	
	Ala	Lys	Tyr	Ile	Lys	Ala	Val	Met	Asn	Ile	Gly	Gly	Pro	Phe	Leu	Gly
				180					185					190		
30	Val	Pro	Lys	Ala	Val	Ala	Gly	Leu	Phe	Ser	Ala	Glu	Ala	Lys	Asp	Met
			195					200					205			
	Arg	Met	Thr	Arg	Thr	Trp	Asp	Ser	Thr	Met	Ser	Met	Leu	Pro	Lys	Gly
35							215						220			
	Gly	Asp	Thr	Ile	Trp	Gly	Gly	Leu	Asp	Trp	Ser	Pro	Glu	Leu	Pro	Asn
	225					230					235					240
40	Ala	Pro	Glu	Met	Glu	Ile	Tyr	Ser	Leu	Tyr	Gly	Val	Gly	Ile	Pro	Thr
					245					250					255	
	Glu	Arg	Ala	Tyr	Val	Tyr	Lys	Leu	Asn	Gln	Ser	Pro	Asp	Ser	Cys	Ile
				260					265					270		
45	Pro	Phe	Gln	Ile	Phe	Thr	Ser	Ala	His	Glu	Glu	Asp	Glu	Asp	Ser	Cys
			275					280					285			
	Leu	Lys	Ala	Gly	Val	Tyr	Asn	Val	Asp	Gly	Asp	Glu	Thr	Val	Pro	Val
50			290				295					300				
	Leu	Ser	Ala	Gly	Tyr	Met	Cys	Ala	Lys	Ala	Trp	Arg	Gly			

	355		360		365	
	Ala Ala Gly Gly Asn Gly Ser Asp Ile Gly His Asp Gln Val His Ser					
	370		375		380	
5	Gly Ile Phe Glu Trp					
	385					
10	<210> 4a					
	<211> 1986					
	<212> DNA					
	<213> Saccharomyces cerevisiae					
15	<220>					
	<221> CDS					
	<222> (1)..(1983)					
	<400> 4					
20	atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct					48
	Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser					
	1 5 10 15					
25	gat gaa aac aat aaa ggg ggt tct gtt cat aac aag cga gag agc aga					96
	Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg					
	20 30					
30	aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt					144
	Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly					
	35 40 45					
35	att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg					192
	Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg					
	50 55 60					
40	aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg					240
	Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu					
	65 70 75 80					
45	att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt					288
	Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe					
	85 90 95					
50	ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt					336
	Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe					
	100 105 110					
55	gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtt					384
	Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val					
	115 120 125					
60	ctc cca caa ggt ata agt tcg ttt att gat gat att cag gct ggt aac					432
	Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn					
	130 135 140					
65	tac tcc aca tct tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt					480
	Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly					
	145 150 155 160					
70	aaa caa ctc tta cgt gat tat aat atc gag gcc aaa cat cct gtt gta					528
	Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val					

	165	170	175	
5	atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile 180 185 190	576		
10	gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp 195 200 205	624		
15	gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp 210 215 220	672		
20	ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn 225 230 235 240	720		
25	ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile 245 250 255	768		
30	gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile 260 265 270	816		
35	ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu 275 280 285	864		
40	gca tat tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys 290 295 300	912		
45	gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu 305 310 315 320	960		
50	att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp 325 330 335	1008		
55	gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val 340 345 350	1056		
60	aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly 355 360 365	1104		
65	gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr 370 375 380	1152		
70	att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser 385 390 395 400	1200		
75	aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser 405 410 415	1248		

	atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct	1296
	Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser	
	420 425 430	
5	tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att	1344
	Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile	
	435 440 445	
10	cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg	1392
	Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met	
	450 455 460	
15	aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa	1440
	Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln	
	465 470 475 480	
20	aga aga gta cat gag cag tac tcg ttc ggc tat tcc aag aat gaa gaa	1488
	Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu	
	485 490 495	
	gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg	1536
	Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met	
	500 505 510	
25	gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac	1584
	Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr	
	515 520 525	
30	ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat	1632
	Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp	
	530 535 540	
35	gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct	1680
	Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro	
	545 550 555 560	
40	gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca	1728
	Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser	
	565 570 575	
	atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga	1776
	Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly	
	580 585 590	
45	att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat	1824
	Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp	
	595 600 605	
50	ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc	1872
	Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser	
	610 615 620	
55	gcg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat	1920
	Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp	
	625 630 635 640	
60	ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag	1968
	Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln	
	645 650 655	

1986

5

10

15

20

25

30

35

40

45

50

55

60

Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
260 265 270

	Gly	Tyr	Glu	Pro	Asn	Lys	Met	Thr	Ser	Ala	Ala	Tyr	Asp	Trp	Arg	Leu
			275					280						285		
5	Ala	Tyr	Leu	Asp	Leu	Glu	Arg	Arg	Asp	Arg	Tyr	Phe	Thr	Lys	Leu	Lys
	290						295					300				
	Glu	Gln	Ile	Glu	Leu	Phe	His	Gln	Leu	Ser	Gly	Glu	Lys	Val	Cys	Leu
	305					310					315					320
10	Ile	Gly	His	Ser	Met	Gly	Ser	Gln	Ile	Ile	Phe	Tyr	Phe	Met	Lys	Trp
					325					330					335	
	Val	Glu	Ala	Glu	Gly	Pro	Leu	Tyr	Gly	Asn	Gly	Gly	Arg	Gly	Trp	Val
15				340					345					350		
	Asn	Glu	His	Ile	Asp	Ser	Phe	Ile	Asn	Ala	Ala	Gly	Thr	Leu	Leu	Gly
			355					360					365			
20	Ala	Pro	Lys	Ala	Val	Pro	Ala	Leu	Ile	Ser	Gly	Glu	Met	Lys	Asp	Thr
		370					375						380			
	Ile	Gln	Leu	Asn	Thr	Leu	Ala	Met	Tyr	Gly	Leu	Glu	Lys	Phe	Phe	Ser
	385					390					395					400
25	Arg	Ile	Glu	Arg	Val	Lys	Met	Leu	Gln	Thr	Trp	Gly	Gly	Ile	Pro	Ser
					405					410					415	
	Met	Leu	Pro	Lys	Gly	Glu	Glu	Val	Ile	Trp	Gly	Asp	Met	Lys	Ser	Ser
30				420					425					430		
	Ser	Glu	Asp	Ala	Leu	Asn	Asn	Asn	Thr	Asp	Thr	Tyr	Gly	Asn	Phe	Ile
			435					440					445			
35	Arg	Phe	Glu	Arg	Asn	Thr	Ser	Asp	Ala	Phe	Asn	Lys	Asn	Leu	Thr	Met
		450					455						460			
	Lys	Asp	Ala	Ile	Asn	Met	Thr	Leu	Ser	Ile	Ser	Pro	Glu	Trp	Leu	Gln
	465					470					475					480
40	Arg	Arg	Val	His	Glu	Gln	Tyr	Ser	Phe	Gly	Tyr	Ser	Lys	Asn	Glu	Glu
					485					490					495	
	Glu	Leu	Arg	Lys	Asn	Glu	Leu	His	His	Lys	His	Trp	Ser	Asn	Pro	Met
45				500					505					510		
	Glu	Val	Pro	Leu	Pro	Glu	Ala	Pro	His	Met	Lys	Ile	Tyr	Cys	Ile	Tyr
			515					520					525			
50	Gly	Val	Asn	Asn	Pro	Thr	Glu	Arg	Ala	Tyr	Val	Tyr	Lys	Glu	Glu	Asp
		530					535					540				
	Asp	Ser	Ser	Ala	Leu	Asn	Leu	Thr	Ile	Asp	Tyr	Glu	Ser	Lys	Gln	Pro
	545					550					555					560
55	Val	Phe	Leu	Thr	Glu	Gly	Asp	Gly	Thr	Val	Pro	Leu	Val	Ala	His	Ser
					565					570					575	
	Met	Cys	His	Lys	Trp	Ala	Gln	Gly	Ala	Ser	Pro	Tyr	Asn	Pro	Ala	Gly
60				580					585					590		

~~et Lys Hi  
00  
la Glu Hi  
eu Lys II  
er Asn L  
6~~

5

625

10

Met

15

[illegible]



## SEQUENCE LISTING

<110> Stymne Dr., Sten  
<120>

<130>

<140>

<141>

<160>

<210> 1

<211> 1986

<212> genomic DNA

<213> *Saccharomyces cerevisiae*

<220>

<221> CDS

<222> (1) .. (1983)

<400> 1

atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct 48  
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser  
1 5 10 15

gat gaa aac aat aaa ggg ggt tct ggt cat aac aag cga gag agc aga 96  
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
20 25 30

aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt 144  
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
35 40 45

att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg 192  
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  
50 55 60

aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg 240  
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
65 70 75 80

att tcc att ctt ggt gca ttc tta ggt gta ctt ttg cag ttt agc ttt 288  
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
85 90 95

ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt 336  
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
100 105 110

00537740-030000

gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtt	384
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val	
115 120 125	
ctc cca caa ggt ata agt tgg ttt att gat gat att cag gct ggt aac	432
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn	
130 135 140	
tac tcc aca tct tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt	480
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly	
145 150 155 160	
aaa caa ctc tta cgt gat tat aat acc gag gcc aaa cat cct gcc gta	520
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val	
165 170 175	
atg gct cct ggt gtc att tct agt gga att gaa agc tgg gga gtt att	576
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile	
180 185 190	
gga gac gat gag tgc gat agt tct ggc cat ttt cgt aaa cgg ctg tgg	624
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp	
195 200 205	
gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg	672
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp	
210 215 220	
ctg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac	720
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn	
225 230 235 240	
ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc	768
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile	
245 250 255	
gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att	816
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile	
260 265 270	
ggc tat gaa ccc aat aaa atg acg agt gct gcc tat gat tgg agg ctt	864
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu	
275 280 285	

gca tat tta gat cta	gaa aga cgc gat agg tac ttt acg aag cta aag	912
Ala Tyr Leu Asp Leu	Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys	
290	295 300	
gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta	960	
Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu		
305 310 315 320		
att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg	1008	
Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp		
325 330 335		
gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt	1056	
Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val		
340 345 350		
aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc	1104	
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly		
355 360 365		
gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc	1152	
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr		
370 375 380		
att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca	1200	
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser		
385 390 395 400		
aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca	1248	
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser		
405 410 415		
atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct	1296	
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser		
420 425 430		
tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att	1344	
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile		
435 440 445		
cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg	1392	
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met		
450 455 460		
aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa	1440	
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln		
465 470 475 480		

aga aga gta cat gag cag tac tgg ttc ggc tat tcc aag aat gaa gaa	1480
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu	
485 490 495	
gag tta aga aaa aat gag cta cac cac aag cac tgg tgg aat cca atg	1536
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met	
500 505 510	
gaa gta cca ctt cca gaa gca ccc cac atg aaa atc tat tgt ata tac	1584
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr	
515 520 525	
ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat	1632
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp	
530 535 540	
gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct	1680
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro	
545 550 555 560	
gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca	1728
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser	
565 570 575	
atg tgt cac aaa cgg gcc cag ggt gct tca ccg tac aac cct gcc gga	1776
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly	
580 585 590	
att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat	1824
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp	
595 600 605	
ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc	1872
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser	
610 615 620	
scg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat	1920
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp	
625 630 635 640	
ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag	1968
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln	
645 650 655	

1986

atg ccc ttc cca atg taa  
Met Pro Phe Pro Met  
660

000000 01/22/86

<211> 661

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 2

Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser  
1 5 10 15  
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
20 25 30  
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
35 40 45  
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  
50 55 60  
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
65 70 75 80  
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
85 90 95  
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
100 105 110  
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
115 120 125  
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
130 135 140  
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
145 150 155 160  
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
165 170 175  
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
180 185 190  
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp  
195 200 205  
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp  
210 215 220  
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn  
225 230 235 240  
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile  
245 250 255  
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
260 265 270  
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
275 280 285  
Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  
290 295 300  
Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu  
305 310 315 320

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp  
 325 330 335  
 Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  
 340 345 350  
 Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly  
 355 360 365  
 Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr  
 370 375 380  
 Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser  
 385 390 395 400  
 Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser  
 405 410 415  
 Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser  
 420 425 430  
 Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile  
 435 440 445  
 Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met  
 450 455 460  
 Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  
 465 470 475 480  
 Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu  
 485 490 495  
 Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met  
 500 505 510  
 Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  
 Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  
 530 535 540  
 Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro  
 545 550 555 560  
 Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser  
 565 570 575  
 Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  
 580 585 590  
 Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
 595 600 605  
 Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
 610 615 620  
 Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
 625 630 635 640  
 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
 645 650 655  
 Met Pro Phe Pro Met  
 660

<210> 1b  
 <211> 2312  
 <212> genomic DNA  
 <213> Schizosaccharomyces pombe  
 <400> 3

ATGGCGTCTT CCAAGAAGAG CAAAACTCAT AAGAAAAAGA AAGAAGTCAA	50
ATCTCCTATC GACTTACCAA ATTCAAAGAA ACCAACTCGC GCTTTGAGTG	100
AGCAACCTTC AGCGTCCGAA ACACAATCTG TTTCAAATAA ATCAAGAAAA	150
TCTAAATTTG GAAAAAGATT GAATTTTATA TTGGGCGCTA TTTTGGAAT	200
ATGCGGTGCT TTTTTTTTCG CTGTTGGAGA CGACAATGCT GTTTTCGACC	250
CTGCTACGTT AGATAAATTT GGGAAATATGC TAGGCTCTTC AGACTTGTTT	300
GATCACATTA AAGGATATTT ATCTTATAAT GTGTTTAAGG ATGCACCTTT	350
TACTACGGAC AAGCCTTCGC AGTCTCCTAG CGGAAATGAA GTTCAAGTTG	400
GTCTTGATAT GTACAATGAG GGATATCGAA GTGACCATCC TGTATTATG	450
GTTCCTGGTG TTATCAGCTC AGGATTAGAA AGTTGGTCGT TTAATAATTG	500
CTCGATTCTT TACTTTAGGA AACGTCTTTG GGGTAGCTGG TCTATGCTGA	550
AGGCAATGTT CCTTGACAAG CAATGCTGGC TTGAACATTT AATGCTTGAT	600
AAAAAACC GCTTGGATCC GAAGGGAATT AAGCTGCGAG CAGCTCAGGG	650
GTTTGAAGCA GCTGATTTTT TTATCACGGG CTATTGGATT TGGAGTAAAG	700
TAATTGAAAA CCTTGCTGCA ATTGGTTATG AGCCTAATAA CATGTTAAGT	750
GCTTCTTACG ATTGGCGGTT ATCATATGCA AATTAGAGG AACGTGATAA	800
ATATTTTTC AAGTTAAAAA TGTTCAATTGA GTACAGCAAC ATTGTACATA	850
AGAAAAAGGT AGTOTTGATT TCTCACTCCA TGGGTTTACA GGTTACGTAC	900
TATTTTTTTA AGTGGGTTGA AGCTGAGGCG TACGGAAATG GTGGACCGAC	950
TTGGGTTAAT GATCATATTG AAGCATTTAT AAATGTGAGT CTCGATGGTT	1000
GTTTGACTAC GTTTCTAACT TTTGAATAGA TATCGGGATC TTTGATTGGA	1050
CCACCCAAAA CAGTGGCAGC GCTTTTATCG GGTGAATGA AAGATACAGG	1100
TATGTAAAT ACATTAAACA TGTTAATATT TAATTTTTGC TAACCGTTTT	1150
AAGCTCAATT GAATCAGTTT TCGGTCTATG GGTAAGCAAT AAATTGTTGA	1200
GATTGTTTAC TAATTTACTG TTTAGTTTGG AAAAAATTTT TTCCCGTTCT	1250
GAGGTATATT CAAAAATACA AATGTGCTCT ACTTTTTCTA ACTTTTAATA	1300
GAGAGCCATG ATGGTTCGCA CTATGGGAGG AGTTAGTTCT ATGCTTCCTA	1350
AAGGAGGCGA TGTTGTATGG GGAAATGCCA GTTGGGTAAG AAATATGTGC	1400

0053740-03000



TGTTAATTTT	TIATTAATAT	TTAGGCTCCA	GATGATCTTA	ATCAAACAAA	1450
TTTTTCCAAT	GGTGCAATTA	TCGATATAG	AGAAGACATT	GATAAGGACC	1500
ACGATGAATT	TGACATAGAT	GATGCATTAC	AATTTTTTAA	AAATGTTACA	1550
GATGACGATT	TTAAAGTCAT	GCTAGCGAAA	AATTATTCCC	ACGGTCTTGC	1600
TTGGACTGAA	AAAGAAGTGT	TAAAAAATAA	CGAAATGCCG	TCTAAATGGA	1650
TAAATCCGCT	AGAAGTAAGA	ACATTAAAGT	TACTAAATTA	TACTAATCCA	1700
AATAGACTAG	TCTTCCTTAT	GCTCCTGATA	TGAAAATTTA	TTGCGTTCAC	1750
CGGGTCGGAA	AACCAACTGA	GAGAGGTTAT	TATTATACTA	ATAATCCTGA	1800
GGGGCAACCT	GTCATTGATT	CCTCGGTAA	TGATGGAACA	AAAGTTGAAA	1850
ATGTGAGAGA	ATTTATGTTT	CAAACATTCT	ATTAAGTGT	TTATTAGGGT	1900
ATTGTTATGG	ATGATGGTGA	TGGAACCTTA	CGAATATTAG	CCCTTGTTT	1950
GGTGTCCAAT	AAAGTTTGGC	AAACAAAAG	GTTTAATCCT	GCTAATACAA	2000
GTATCACAAA	TTATGAAATC	AAGCATGAAC	CTGCTGCGTT	TGATCTGAGA	2050
GGAGGACCTC	GCTCGGCAGA	ACACGTCGAT	ATACTTGGAC	ATTCAGAGCT	2100
AAATGTATGT	TCATTTTACC	TTACAAATTT	CTATTACTAA	CTCTTGAAAT	2150
AAGGAATTA	TTTTAAAAGT	TTCATCAGGC	CATGGTGAAT	CGGTACCAAA	2200
CCGTTATATA	TCAGATATCC	AGTACGGACA	TAAGTTTGT	AGATTGCAAT	2250
TACTAACTA	ACCGAACAGG	GAAATAATAA	ATGAGATAAA	TCTCGATAAA	2300
CCTAGAAATT	AA				2312

ATGCCCTTA	TTCATCGGAA	AAAGCCGACG	GAGAAACCAT	CGACGCCGCC	50
ATCTGAAGAG	GTGGTGCACG	ATGAGGATTC	GCAAAAGAAA	CCACACGAAT	100
CTTCCAAATC	CCACCATAAG	AAATCGAACG	GAGGAGGCAA	GTGGTCGTGC	150
ATCGATTCTT	GTGTGTGGTT	CATTGGGTGT	GTGTGTGTAA	CCTGGTGGTT	200
TCTTCTCTTC	CTTTACAACG	CAATGCTGTC	GAGCTTCCCT	CAGTATGTAA	250
CGGAGCGAAT	CACGGGTCCT	TTGCCTGACC	CGCCCCGGTG	TAAGCTCAAA	300
AAAGAAGGTC	TTAAGGCGAA	ACATCCTGTT	GTCTTCATTC	CTGGGATTGT	350
CACCGGTGGG	CTCGAGCTTT	GGGAAGGCTA	ACAATGCGCT	GATGGTTTAT	400
TTAGAAAACG	TTTGTGGCGT	GGAACCTTTG	GTGAAGTCTA	CAAAAGGTGA	450
GCTCAACAAT	TCTCACTCTT	CCTTTATATT	GGGATTGGA	TTGGATCTGA	500
TGAGATCAGC	CACCTGTGTC	TTCTTCAACA	TCACTCAAAC	TTTAACTCCA	550
TGTTTGICTG	TCTTACTCTT	TACTTTTTTT	TTTTTTTGAT	GTGAACGCT	600
ATTTTCTTAA	GAGACTATTT	CTGTATGTGT	AAGGTAAGCG	TTCCAAGGAC	650
GTAATGGCT	TGGACTATTT	CTGTTTGATT	GTAACTTTA	GGATATAAAA	700
TAGCTGCCCT	GGAATTTCAA	GTCATCTTAT	TGCCAAATCT	GTTGCTAGAC	750
ATGCCCTAGA	GTCCGTTTCT	AACAAGTTAC	TTCTTTTACT	GTCGTTGCGT	800
GTAGATTAG	CTTGTGTAG	CGTATAATGA	AGTAGTGTTT	TATGTTTTGT	850
TGGGAATAGA	GAAAGTTCTAA	CTACATCTGT	GGAAAGTGTC	TTCAGGCTGT	900
GATAGAGGAC	TGTTGCTTTA	TTATTCAACT	ATGTATATGT	GTAATTAAG	950
CTAGTTCCTT	TTTGATCTTT	CAGCTCAATG	TGCTTTTCTC	AATTTTTTTC	1000
TCAATTTCAA	AGTTTCACAT	CGAGTTTATT	CACATGTCTT	GAATTCGTC	1050
CATCCTCGTT	CTGTTATCCA	GCTTTGAACT	CCTCCCGACC	CTGCTATGGA	1100
TATATTAAAA	AAAAGTGTT	TTGTGGGTG	CATCTTTGTT	ACGATCTGCA	1150
TCTTCTTCTT	TCGGCTCAGT	GTTTATGTTT	TTGCTATGGT	AGAGATGGGC	1200
AATGTTATTG	TTGATGGTAA	CAGTGGTATA	GTTGATAGTA	TCTTAACTAA	1250
TCAATTATCT	CTTTGATCA	GGCCTCTATG	TTGGGTGGAA	CACATGTCAC	1300
TTGACAATGA	AAGTGGGTG	GATCCAGCTG	GTATTAGAGT	TCGAGCTGTA	1350

TCAGGACTCG TGGCTGCTGA CTACTTTGCT CCTGGCTACT TTGTCTGGGC	1400
AGTGCTGATT GCTAACCTTG CACATATTGG ATATGAAGAG AAAAATATGT	1450
ACATGGCTGC ATATGACTGG CGGCTTTTCTG TTCAGAACAC AGAGGTTCTT	1500
TTCTCATCGT TCTTTCTATT ATTCTGTTCC ATGTTACGTT TCTTTCTTCA	1550
TTACTTAAGG CTTAAATATG TTTTATGTTG AATTAAATAGG TACGTGATCA	1600
GACTCTTAGC CGTATGAAAA GTAAATATAGA GTTGATGGTT TCTACCAACG	1650
GTGGAAAAAA AGCAGTTATA GTTCCGCTATT CCATGGGGGT CTGTATTTT	1700
CTACATTTTA TGAAGTGGGT TGAGGCACCA GCTCCTCTGG GTGGCGGGGG	1750
TGGGCCAGAT TGGTGTGCAA AGTATATATA GGCGGTGATG AACATTGGTG	1800
GACCATTTCT TGGTGTTCCT AAAGCTGTTG CAGGGCTTTT CTCTGCTGAA	1850
GCAAAGGATG TTGCAATTGC CAGGTATTGA ATATCTGCTT ATACTTTTGA	1900
TGATCACAAC CTGGCTCTG GAACTCAAAG TTATTCTACT AAATATCAAT	1950
TCTAATAACA TTGCTATATT ATCGCTGCAA CTGACATTGG TTGATTATTT	2000
TTGCTGCTTA TGTAAGTGAA ACTCTCTTGA GATTAGACAA ATGATGAATT	2050
GATAATTCTT ACGCATTGCT CTGTGATGAC CAGTTTCTTA GCTTCGACGA	2100
TAACATTGT CATACTGTCT TTTGGAGGGC ATTGAATTTT GCTATGGAAA	2150
GCGCTGGAGC TTCCATGCTT GCATTCTTTA CCAATTAGCG TTATTCTGCT	2200
TCTTTCAATT TTCTTGATA TGCATCTATG GTCTTTTATT TCTTCTTAAT	2250
TAAAGACTCG TTGGATTAGT TGCTCTATTA GTCATTGGT TCCTTAATAT	2300
AGAACTTTAC TTTCTTCGAA AATTGCAGAG CGATTGCCCC AGGATTCTTA	2350
GACACCGATA TATTAGACT TCAGACCTTG CAGCATGTAA TGAGAATGAC	2400
ACGCACATGG GACTCAACAA TGTCTATGTT ACCGAAGGGA GGTGACACGA	2450
TATGGGGCGG GCTTGATTGG TCACCGGAGA AAGGCCACAC CTGTTGTGGG	2500
AAAAGCAAA AGAACAACGA AACTTGTGGT GAAGCAGGTG AAAACGGAGT	2550
TTCCAAGAAA AGTCCTGTTA ACTATGGAAG GATGATATCT TTTGGGAAAG	2600
AAGTAGCAGA GGCTGCGCCA TCTGAGATTA ATAATATTGA TTTTCGAGTA	2650
AGGACATATA AATCATAATA AACCTTGTAC ATTTTGTGAT TGTATGATGA	2700
ATATCTGTAC ATTTTATCTG GTGAAGGGTG CTGTCAAAGG TCAGAGTATC	2750
CCAATCACA CCTGTCTGTA CGTGTGGACA GAGTACCATG ACATGGGAAT	2800
TGCTGGGATC AAAGCTATCG CTGAGTATAA GGTCTACACT GCTGCTGAAG	2850

CTATAGATCT ACTACATTAT GTTGCTCCTA AGATGATGGC GCGTGGTGCC	2900
GCTCATTTCI CTTATGGAAT TGCTGATGAT TTGGATGACA CCAAGTATCA	2950
AGATCCCAAA TACTGGTCAA ATCCGTTAGA GACAAAGTAA GTGATTTCTT	3000
GATTCCAACI GIATCCTTCG TCCTGATGCA TTATCAGTCT TTTTGTITTC	3050
GGTCTTGTTG GATAIGGTTT TCAGCTCAA GCTTACAAAG CTGTTTCTGA	3100
GCCTTTCTCA AAAAGGCTTG CTCAGTAAAT TTGAGGTGCT AAAGTTGATA	3150
CATGTGACTC TTGCTTATAA ATCCTCCGTT TGGTTTGTTC TGCTTTTTC	3200
GATTACCGAA TGCTCCTGAG ATGGAAATCT ACTCATTATA CGGAGTGGGG	3250
ATACCAACGG AACGAGCATA CGTATACAAG CTTAACCAGT CTCCCGACAG	3300
TTGCATCCCC TTTCAGATAT TCACTTCTGC TCACGAGGAG GACGAAGATA	3350
GCTGTCTGAA AGCAGGAGTT TACAATGTGG ATGCGGATGA AACAGTACCC	3400
GTCCTAAGTG CCGGGTACAT GTGTGCAAAA GCGTGGCGTG GCAAGACAAG	3450
ATTCAPCCCT TCCGGAATCA AACTTTATAT AAGAGAATAC AATCACTCTC	3500
CGCCGGCTAA CCTGTTGGAA GGGCGCGGGA CGCAGAGTGG TGCCCATGTT	3550
GATAICATGG GAACTTTTGC TTTGATCGAA GATATCATGA GGGTTGCCGC	3600
CGGAGGTAAC GGGTCTGATA TAGGACATGA CCAGGTCCAC TCTGGCATAT	3650
TGAATGGTC GGAGCGTATT GACCTGAAGC TGTGA	3685

<210> 5  
 <211> 402  
 <212> cDNA  
 <213> Arabidopsis thaliana  
 <220>  
 <221> CDS  
 <222> (120) .. (402)  
 <400> 5

AGAAACAGCTCTTTGTCTCT CTCGACTGATCTAACAATCC CTAATCTGTGTCTCTAAATTC 60  
 CTGGACGAGATTTGACAAAG TCCGTATAGCCTAACCTGGT TTAATTTCAAGTGACAGAT 119  
 ATG CCC CTT ATT CAT CGG AAA AAG CCG ACG CAG AAA CCA TCG ACG CCG 167  
 Met Pro Leu Ile His Arg Lys Lys Pro Thr Glu Lys Pro Ser Thr Pro  
 CCA TCT GAA GAG GTG GTG CAC GAT GAG GAT TCG CAA AAG AAA CCA CAC 215  
 Pro Ser Glu Glu Val Val His Asp Glu Asp Ser Gln Lys Lys Pro His  
 GAA TCT TCC AAA TCC CAC CAT AAG NAA TCG AAC GGA GGA GGG AAG TGG 263  
 Glu Ser Ser Lys Ser His His Lys ??? Ser Asn Gly Gly Gly Lys Trp  
 TCG TGC ATC GAT TCT TGT TGT TGG TTC ATT GGG TGT GTG TGT GTA ACC 311  
 Ser Cys Ile Asp Ser Cys Cys Trp Phe Ile Gly Cys Val Cys Val Thr  
 TGG TGG TTT CTT CTC TTC CTT TAC AAC GCA ATG CCT GCG AGC TTC CCT 359  
 Trp Trp Phe Leu Leu Phe Leu Tyr Asn Ala Met Pro Ala Ser Phe Pro  
 CAG TAT GTA ACG GAG CCG AAT CAC GNG TCC TTT GCC TTA CCC G 402  
 Gln Tyr Val Thr Glu Pro Asn His ??? Ser Phe Ala Leu Pro

00537710-033000

[illegible]

<210> 7<sup>h</sup>  
 <211> 115  
 <212> PRT  
 <213> Zea mays  
 <400> 7

Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly  
 1 5 10 15

Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly  
 20 25 30

Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val  
 35 40 45

Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly  
 50 55 60

Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile  
 65 70 75 80

Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly  
 85 90 95

Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys  
 100 105 110

Leu Lys Leu  
 115

000000-0722500

<210> 8 *le*  
 <211> 616  
 <212> cDNA  
 <213> *Neurospora crassa*  
 <400> 8

GGTGGCGAAG ACGANGGCGG AAGTTGGAGG CTAACGAGAA TGACNCTCGG	50
AGATGGATCT ACCCTCTAGA GACACGACCA CCNTTGCACC CAGCCTCAAG	100
GTNTACNGTT TNTATGGGTA GGAAGCCGAG GGAGCGAGCC TACATCTATC	150
TGGCGCCCCG TCCCGGGACG ACAACGCATC TTTAGATGAC GATCGATACG	200
ACTTTGACTN AGGGGCACAT TGACCACGGT GTGATTTTGG GCGAAGGCGA	250
TGGCACAGTG AACCTTATGA GTTTGGGGTA CCTGTGCAAT AAGGGGTGGA	300
AAATGAAGAG ATACAATCCT GCGGGCTCAA AAATAACCGT GGTCCGAGATG	350
CCGCATGAAC CAGAACGGTT CAATCCGAGA GGAGGGCCGA ATACGGCCGA	500
CTTAAATATG TAGAAAAGGT TGAAATTTAT GAAGAGTAAT TAAATACGGC	550
ACATAGGTTA CTCAATAGTA TGAATAATTA AAAAAAATT TTTTTCTAA	600
AAAAAAAAAA AAAAAA	616

00537710-033000



[illegible]

TTAACAGAGG AGCTGATGAC TCCGGGAGTG CCAGTCACTT GCATATATGG 1300  
GAGAGGAGTT GATACACCGG AGGTTTTGAT GTATGGAAAA GGAGGATTCG 1350  
ATAAGCAACC AGAGATTAAG TATGGAGATG GAGATGGGAC GGTTAATTG 1400  
GCGAGCTTAG CAGCTTTGAA AGTCGATAGC TTGAACACCG TAGAGATTGA 1450  
TGGAGTTTCG CATACTCTA TACTTAAAGA CGAGATCGCA CTTAAAGAGA 1500  
TTATGAAGCA GATTTCATT ATTAATTATG AATTAGCCAA TGTTAATGCC 1550  
GTCAATGAAT GA 1562

000000-0722560

<210> 10  
 <211> 3896  
 <212> genomic DNA  
 <213> Arabidopsis thaliana  
 <400> 10

ATGGGAGCGA	ATTCGAAATC	AGTAACGGCT	TCCTTCACCG	TCATCGCCGT	50
TTTTTCTTG	ATTTGCGGTG	GCCGAACCTG	GGTGGAGGAT	GAGACCGAGT	100
TTCACGGCGA	CTACTCGAAG	CTATCGGCTA	TAATCATTCG	CCCAATTCCG	150
TCGACGCAGC	TACGAGCGTG	GTGATCCTT	GACTGTCCAT	ACACTCCGTT	200
CGACTTCAAT	CCGCTCGACC	TCGTATGGCT	AGACACCACT	AAGGTCCGTG	250
ATCTTCATTT	CCCTCGCTCC	TTATTCTGTC	GGTCGAGTCA	CTTGTTGATG	300
AATTCCAAGC	GAAATATAGC	AATGAAGCAT	GTCTCGTCTC	TCATTATTGAT	350
TCGTTCAATG	GTCAACAGTG	ACGCTTCTGA	ATCTGAGTTT	AGAGTCATAT	400
AAACAGCTG	ACTCGGCGAG	TCTTTCCCAT	CGCTTTTGGT	TCGCTAAATG	450
TAGCGCAATG	AATGTGTAAT	TAGTCTGCGC	TTTTTATTCA	ACTAGATCTG	500
CAAGTTTTTC	AGAGTGCTCA	ATAGTAGTTA	GAAATGTGTA	GGTCATTTTA	550
CTTGTGCATT	GTGATTCTTT	TGGTGTGTTG	TTACTGATCG	ACGTGATGGA	600
TGGTTTACAG	CTTCTTTCTG	CTGTCAACTG	CTGGTTTAA	TGTATGGTGC	650
TAGATCCTTA	TAATCAAACA	GACCATCCCG	AGCTGAAGTC	ACGGCCTGAC	700
AGTGGTCTTT	CAGCCATCAC	AGAAATGGAT	CCAGGTTACA	TAACAGGTAG	750
TTTCGGATTT	TTCTTTCTTT	TGAGTTTCT	TCAATTTGAT	ATCATCTTGT	800
TGTGATATAA	TATGGCTAAG	TTCATTGATT	TGGTCAATTT	TCAGGTCCTC	850
TTTCTACTGT	CTGGAAGAG	TGGCTTAAGT	GGTGTGTTGA	GTTTGGTATA	900
GAAGCANATG	CAATTGTGCG	TGTTCCATAC	GATTGGAGAT	TGTCACCAAC	950
CAAAATGGAA	GAGCGTGACC	TTTACTTTCA	CAAGCTCAAG	TTAGTCCTTA	1000
TCAGGCTAAT	GTCTTTTATC	TTCTCTTTT	ATGTAAGATA	AGCTAAGAGC	1050
TCTGGTCGTC	TTCTTTTGTG	CAGGTTGACC	TTTGAAGCTG	CTTTAAAGCT	1100
CCGTGGCGGC	CCTTCTATAG	TATTTGCCCA	TTCAATGGGT	AATAATGTCT	1150
TCAGATACTT	TCTGGAATGG	CTGAGGCTAG	AAATTGCACC	AAACATTAT	1200
TTGAATGCGC	TTGATCAGCA	TATCCATGCT	TATTTCCGCT	TGGTACCGG	1250
CCTACTATCC	TAAAGTTACC	ATTTTATTTT	TCTCTAATT	GGGGGAGTTA	1300
CTTGTGACT	TACTGGATTG	AGCTCGATAC	CTGATTTGTT	GTTGATTTAG	1350
GAGCTCCTCT	TCTTGGTTCT	GTTGAGGCA	TCAATCTAC	TCTCTCTGGT	1400
GTAACGTTTG	GCCTTCCTGT	TTCTGAGGTG	ACCTCTGACT	TCTCTTTAGT	1450
TTTAAGTAGT	TGATATCAAC	CAGGTCTTAT	AAGTCACTGG	ATTTTCCTTT	1500
TGAAGTATT	ACTTTTGTTA	ATTGAAGTGC	TGTACGCGAT	ATGGTATCTG	1550
TAGATCTTGA	AGTGCTAGTT	ATCAAGAAGC	ATATTGTGGG	TAGTATACCT	1600
GTCAGCGGCC	TTAGCTAATA	CAACCAAACC	ACATGTACAC	TGATTTAGTT	1650
TTCAAGATTAT	TATGGTAGAC	TTTAAGTTGA	CAAGAAACTT	TGACTGAAAT	1700
CTTTTATTTT	TAATAGGCTA	TGATTTGTTT	ATTGAATCA	TGTGACATAT	1750
TGACATGCGC	TTCTCATGTT	TTTTGTTGGC	AAGGCTTCAG	GGAAGTGCCT	1800
GGTGTGTTG	CAATTCCTTT	GCGTCGTGAT	TGTGGCTTAT	GCCATTTTCA	1850
AAGAATTGCA	AGGGTGATAA	CACATTCTGG	ACGCATTTT	CTGGGGGTGC	1900
TGCNAAAGAAA	GATAAGCGCG	TATACCACTG	TGATGAAGAG	GAATATCAAT	1950
CAAAATATTC	TGGCTGGCCG	ACAAATATTA	TTAACATTGA	AATTCCTTCC	200
ACTAGCGGTT	AGACTCTGTA	TATGCAACTG	TAACACTAAG	AAAAGTTTCA	2050
CCAAGAATGT	TCACTCTCAT	ATTTCCGTTCC	TTTGATGTGT	ATCCATCAGT	2100
TACAGAAACA	GCTCTAGTCA	ACATGACCAG	CATGGAATGT	GGCCTTCCCA	2150
CCCTTTTGTG	TTTCAAGAGC	CGTGAAGTAG	CAGATGGGAC	TCTTTTCAAA	2200
GCAATAGAAAG	ACTATGACCC	AGATAGCAAG	AGGATGTTAC	ACCAGTTAAA	2250
GAAGTACGTA	CCTTCTTTTG	TGATAAGAAA	TATTGCTCAT	CGATCATCAC	2300
TTGCTGGCTT	CTTGTACGTC	AAATTGTTTT	GTTTAAATCT	CTATATCAAT	2350
TGTTTATATG	CTTTGTCTTT	CTTACTATAA	GAAACAAGTA	TAATCAGAAA	2400
CCTTATTATT	GATTATCAGT	TCTCTCCTTA	TATTATGGAA	TGCTTTTTC	2450
GTTTACAGTT	ATGAATGCAA	AAGGGGGTAT	TTAGTTGAT	TGATTCTCTC	2500
ATTCTCTAGT	TTGTTTTGAC	TAATAGCGTC	AATTTTGT	TTCTAGCAAA	2550
TCTTTGTGAA	TTATATATAA	CATGCTAACT	ATACTTTTCA	GGTGTATCA	2600
TGATGACCCT	GTTTTTAATC	CTCTGACTCC	TTGGGAGAGA	CCACTATAA	2650
AAAAATGATT	TTGCATATAT	GGTGCTCATC	TAAAGACAGA	GGTATGATGC	2700
ATTCTCAATA	TCACATTATG	CGTTGACTTT	GTTATTATAT	TCCCATTG	2750

00537710.0330000

GTTTGCAATA	TCTTTTTGAA	TTATGATTTA	TCTTCTCCCT	TGCATCTTAT	2800
CCTATTAAGC	GTTAAAGGTA	CTAAATGTAT	GAAGCTGTCT	GTCTAGGTT	2850
GGTTATTACT	TTGCCCAAG	TGGCAACCT	TATCCTGATA	ATTGGATCAT	2900
CACGGATATC	ATTTATGAAA	CTGAAGGTT	CCTCGTGTCA	AGGTAATTTT	2950
CCGCAATGGC	ACAAGTAAAA	CAGGAAGGCA	AAGTCTTCTG	TATCAGTCTA	3000
GTGGCATGTT	ATCTCAGTTG	CATAAGCAAA	TTATTAAACA	ACTAAAATTT	3050
AAGTACTTTT	TTATCATTCC	TTTGTGACCT	AGTGGATGAT	CAGTGGCTTA	3100
AAGTGGGAAG	AGGTGTTGCA	TGAAACATGA	CACTTGTATC	AAAGATNACT	3150
ACCAAAACAA	AACTAACCCA	TTTCTGATTT	TCATATTATT	AGGAGTAGTC	3200
GTGCTTTTAA	AAAATTTGTT	TTAAGAAACC	GAAAAACTAG	TTCATATCTT	3250
GATTGTGCAA	TATCTGCAGG	TCTGGAACCTG	TGGTTGATGG	GAACGCTGGA	3300
CCTATAACTG	GGGATGAGAC	GGTAAGCTCA	GAAGTTGGTT	TTGAAATTAT	3350
CTTCTTGCAA	ACTACTGAAG	ACTAAGATAA	TACTTGCTTC	TGGAACACTG	3400
CTTGCTATGT	TCTCTAGTAC	ACTGCAATAT	TGACTCTCCG	CTACTTTTAT	3450
TGATTATGAA	ATTGATCTCT	TATAGGTACC	CTATCATTCA	CTCTCTTGGT	3500
GCAAGAATTG	GCTCGGACCT	AAAGTTAACA	TAACAATGGC	TCCCCAGGTA	3550
CTCTTTTTTA	GTTCCCTCACC	TTATATAGAT	CAAACCTTAA	GTGTACTTTT	3600
CTGGTTAIGT	GTTGATTTAC	CTCCAATTTG	TTCTTTCTAA	AAATCATATA	3650
TCCTGTACT	CCTCAAGAAC	TTGTATTAA	CTAAACGAGA	TTCTCATTGG	3700
GAAAATAAAA	CAACAGCCAG	AACACGATGG	AAGCGACGTA	CATGTGGAAC	3750
TAAATGTTGA	TCATGAGCAT	GGGTCAGACA	TGATAGCTAA	CATGACAAAA	3800
GCACCAAGGG	TTAAGTACAT	AACCTTTTAT	GAAGACTCTG	AGAGCATTCC	3850
GGGGAAGAGA	ACCGCAGTCT	GGGAGCTTGA	TAAAAGTGCG	TATTAA	3896

00537710-033000

CTGGGGGCCAA AAGTGAACAT AACAAAGGACA CCACAGTCAG AGCATGATGT	50
TCAGATGTAC AAGTGCATCT AAATATAGAG CATCAACATG GTGAAGATAT	100
CATTCCCAAT ATGACAAAGT TACCTACAAT GAAGTACATA ACCTATTATG	150
AGGATTCTCA AAGTTTTCCA GGGACRAGAA CAGCAGTTTG GGAGCTTGAT	200
AAAGCAAATC ACAGGAACAT TGTCAGATCT CCAGCTTTGA TCGGGGAGCT	250
GTGGCTTGAG ATGTGGCATG ATATTCTCC TGATAAAAAG TCCAAGTTTG	300
TTACAAAAGG TGGTGTCTGA TCCTCACTAT TTTCTTCTAT AAATGTTTGA	350
GTTTGTATTG ACATTGTAAG TATTGCAACA AAAAGCAAAG CGTGGGCCTC	400
TGAGGGATGA GGA CTGCTAT TGGGATTACG GGAAAGCTCG ATGTGCATGG	450
GCTGAACATT GTGAATACAG GTTAGAATAT TCAATTATA TTTTGCAAAA	500
TATTCTCTTT TTGTGTATTT AGGCCACCTT TCCCCGGTCA CAACGATGCA	550
GATATGTATT CGGGGATGTT CACCTGGGAC AGAGTTGCAG ATTGAAGAGT	600
TCTACAICTC ACATCCTGTC ACACTATGTG TGATATTTAA GAAACTTTGT	650
TTGGCGGAAC AACAAAGTTG CACAAACATT TGAAGAAGAA AGCGAAATGA	700
TTTCAGAGAG	709